

Type 2 Diabetes: ***New Insights from Genomics***

Demystifying Medicine - Feb 24, 2009

Lori Bonnycastle, Ph.D.
National Human Genome Research Institute

Risk Factors for Diabetes

- **Genetic**
- **Environmental**

***“If your child looks like you
----its genetics”***

***“If your child looks like your neighbour
----its environment”***

T. Hansen

TOPICS:

- 1. Searching for those darn genes ---GWAs**
- 2. What have we learned?**
- 3. How can we use this information?**
- 4. Next steps.....**

TOPICS:

1. Searching for those darn genes ---GWAs

Familial Clustering

Sibling relative risk (λ_s):

$$\text{T2D } \lambda_s = \frac{\text{risk to siblings}}{\text{population risk}} = \frac{\sim 30\%}{10\%} = \sim 3$$

T1D $\lambda_s = 15$

Crohn's Disease $\lambda_s = 20-35$

Cancer $\lambda_s =$

Breast (EO) ~ 3

Colorectal >4

Ovarian 10-60

T2D Genes - Linkage / Candidate Gene Based

Generally accepted as **true** T2D susceptibility genes:

- 1. *PPAR γ* - Peroxisome Proliferator-Activated Receptor Gamma**
 - Adipocyte differentiation
 - Regulates FA storage and glucose metabolism
- 2. *KCNJ11* - Potassium inwardly rectifying channel KIR 6.2**
 - Islet ATP sensitive potassium channel
 - Role in glucose induce insulin secretion
- 3. *TCF7L2* - Transcription factor 7-like 2**
 - Transcription factor in Wnt signaling pathway --gene activation

Genetic Approaches

1. Linkage analysis

- Family based study ---identify linked regions

2. Candidate gene studies (usually based on perceived biological connection)

3. Genome Wide Association study (GWAs)

- Usually population based

Single Nucleotide Polymorphisms (SNPs)

GAAATAATTAATCTTTTCCTTCCTTCTCCTATTTTGTCTTTACTTCAATTTATTTATTTATTATTAATATTATTATTTTGTGAG
ACGGAGTTT **C/**

ACTCTTGTTGCCAACCTGGAGTGCAGTGGCGTGATCTCAGCTCACTGCACACTCCGCTTTCCTGGTTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGTCACACACCACCACGCCCCGGCTAATTTTTGTATTTTATAGTAGAG
TTGGGGTTTACCATTGTTGGCCAGACTGGTCTG **GAA**CTCTGACCTTGTGATCCGCCAGCCTCTGCCTCCCAAAGAGC
TGGGATTACAGGCGTGAGCCACCGCGCTCG **CCCTTT** CATCAATTTCTACAGCTT **TTTT**CTTTGCCTGGACTTTACA
AGTCTTACCTTGTTCTGC **C/**

TTCAGATATTTGTGTGGTCTCATTCTGGTGTGCCAGTAGCTAAAAATCCATGATTTGCTCTCATCCCACTCCTGTTGTTT
ATCTCCTCTTATCTGGGGTCAC **A/**

CTATCTCTTCTGATTGCATTCTGATCCCCAGTACTTAGCATGTGCGTAACAACCTCTGCCTCTGCTTTCCAGGCTGTT
GATGG **GGTGCT**TTTCATGCCTCAGAAAAATGCATTGTAAGTTAAATTATTAAGATTTTAAATATAGGAAAAAAGTAAGC
AAACATAAGGACAAAAAGGAAAGAACATGTATTCTAATCCATTATTTATTATACAATTAAGAAATTTGGAAACTTTAGAT
TAACTGCTTTTAGAGATGGAGATGTAGTAAGTCTTTTACTCTTTACAAAATACATGTGTTAGCAATTTTGGGAAGAATA
GTAACACCCGAACAGT **G/**

TAATGTGAATATGTCACCTTACTAGAGGAAAGAAGGCACTT **GAAA** AACATCTCTAAACCGTATAAAAAACAATTACATCAT
AATGATGAAAACCCAAGGAATTTTTTTAGAAAACATTAC **CAGGG**CTAATAACAAAGTAGAGCCACATGTCATTTATCTTC
CCTTTGTGTCTGTGTGAGAATTCTAGAGTTATATTTGTACATAGCATGGAAAAATGAGAGGCTAGTTTATCAACTAGTTC
ATTTTAAAGTCTAACACATCCTAGGTATAGGTGAACTGTCCTCCTGCCAATGTATTGCACATTTGTGCCCAGATCCAG
CATAGGGTATGTTTGCCATTTACAAACGTTTATGTCTTAAGAGAGGAAATATGAAGAGCAAAACAGTGCATGCTGGAGA
GAGAAAGCTGATACAAATATAAA **T/**

GAAACAATAATTGGAAAAATTGAGAACTACTCATTTTCTAAATTACTCATGTATTTTCTAGAAATTAAGTCTTTTAATT
TTGATAAATCCCAATGTGAGACAAGATAAGTATTAGTGATGGTATGAGTAATTAATATCTGTTATATAATATTCAATTTTATA
GTGGAAGAAATAAAATAAAGGTTGTGATGATTGTTGATTATTTTTTCTAGAGGGGTTGTCAGGGAAAGAAATTGCTTTTT

Courtesy of T. Manolio

SNPs 1 / 300 bases
~ 10 million across genome

Association Analysis

SNP A

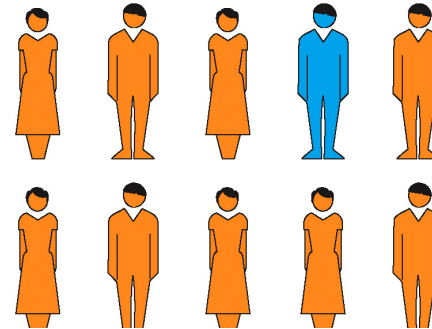


Diabetics

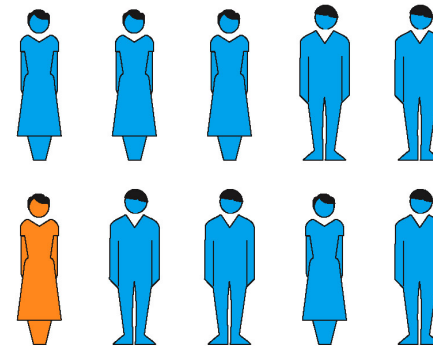


Unaffected

SNP B



Diabetics



Unaffected

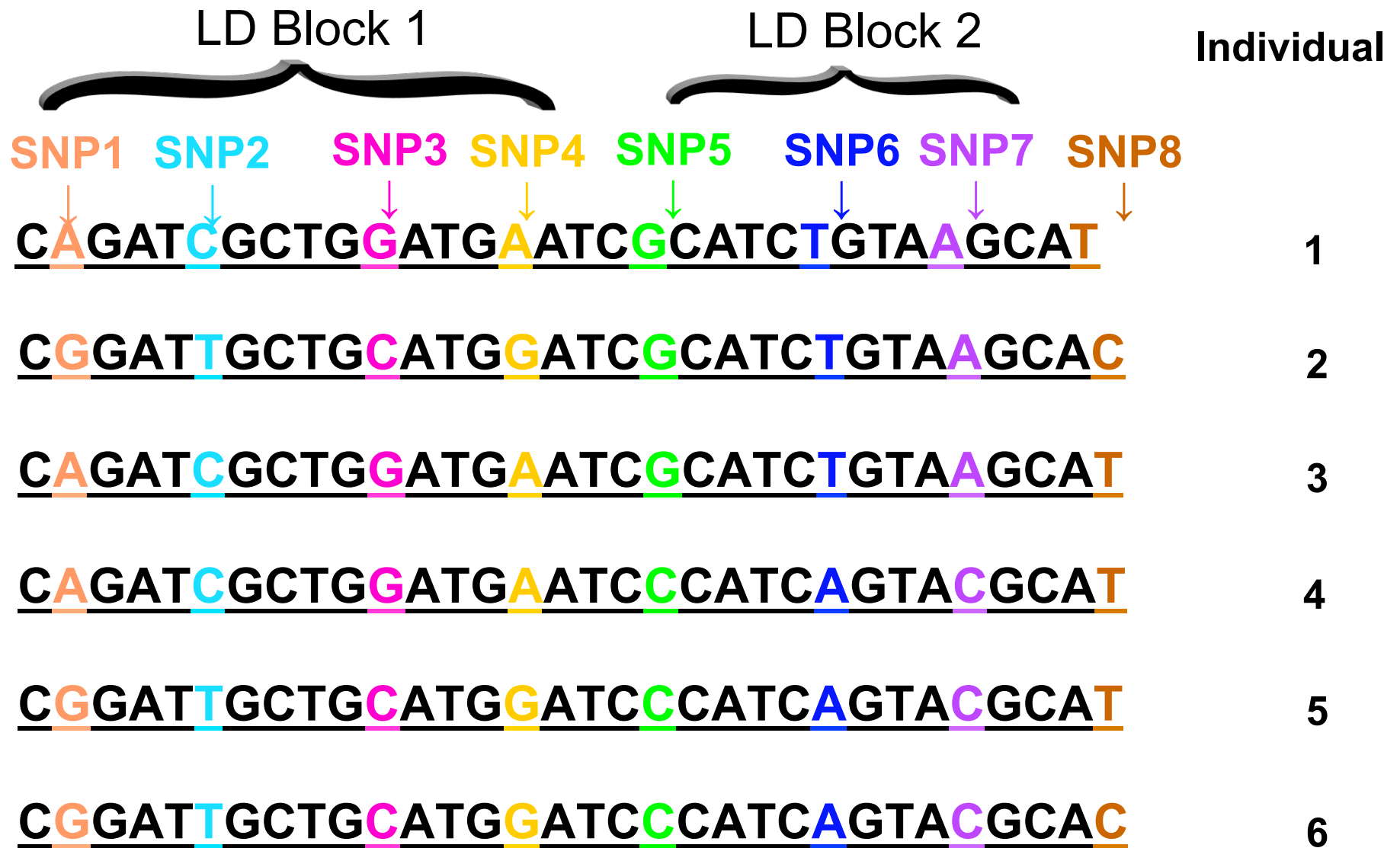
Strength of results measured by p-value

Genome-Wide Association Study

- Aim to interrogating all 10 million variable points across human genome
- Variation inherited in groups, or “LD” (linkage disequilibrium) blocks, such that alleles of nearby SNPs are correlated
- HapMap (Haplotype Map) : A map of common patterns of human genetic variation
- Can use HapMap to determine the
 - pattern of SNPs (“haplotype”) within a locus
 - SNPs can be used to “tag” other correlated SNPs

→ ~80% of common variation covered with only 300-500K SNPs

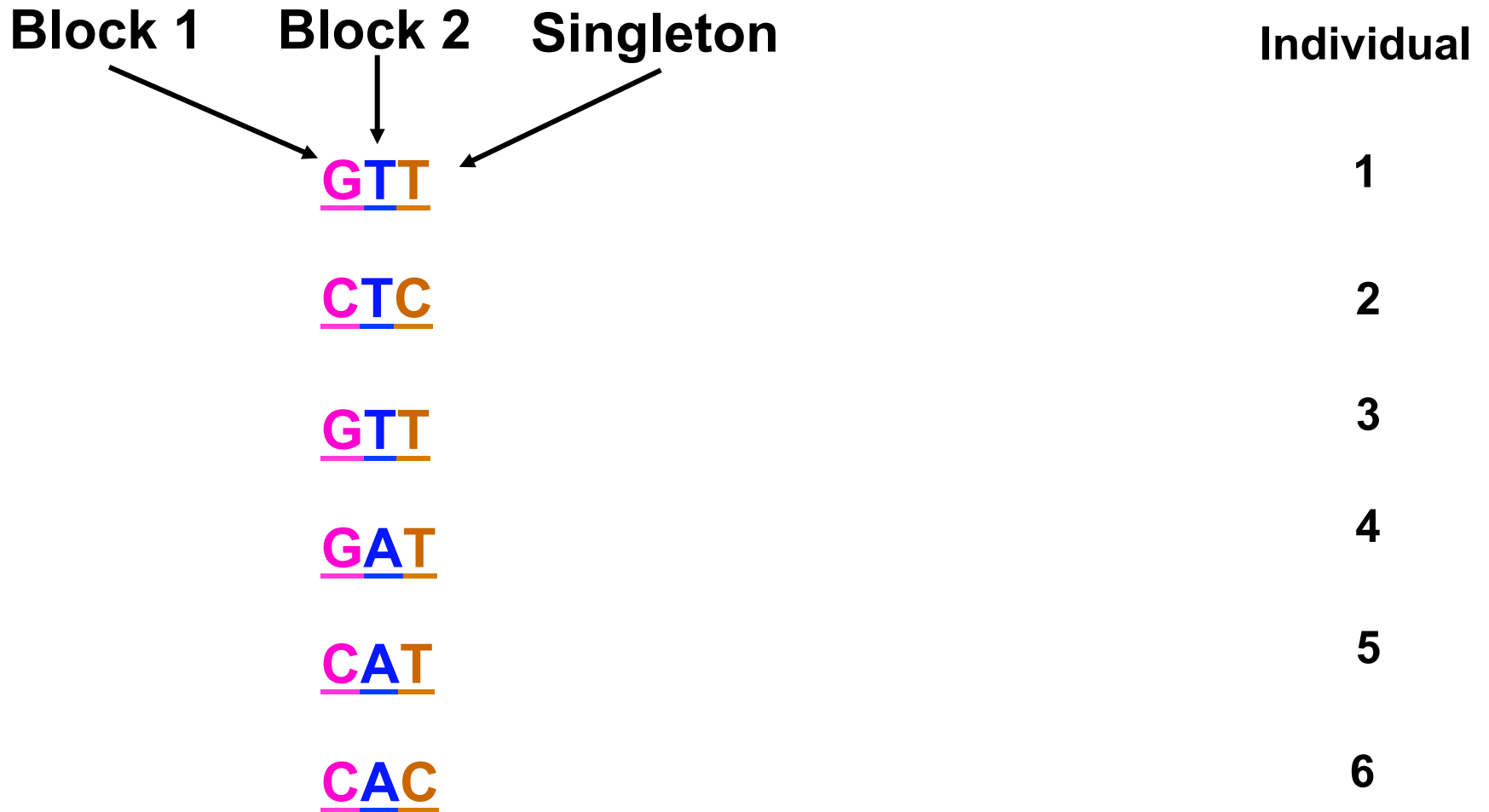
One Tag SNP May Serve as Proxy for Many



One Tag SNP May Serve as Proxy for Many

Block 1		Block 2		Individual
SNP3		SNP6	SNP8	
	<u>G</u>	<u>T</u>	<u>T</u>	1
	<u>C</u>	<u>T</u>	<u>C</u>	2
	<u>G</u>	<u>T</u>	<u>T</u>	3
	<u>G</u>	<u>A</u>	<u>T</u>	4
	<u>C</u>	<u>A</u>	<u>T</u>	5
	<u>C</u>	<u>A</u>	<u>C</u>	6

One Tag SNP May Serve as Proxy for Many



Current T2D Susceptibility Genes ($p < 5 \times 10^{-8}$)

> 32,000 samples

Gene	OR	p-value
<i>TCF7L2</i>	1.37	1.0E-48
<i>IGF2BP2</i>	1.14	8.9E-16
<i>CDKN2A/B</i>	1.20	7.8E-15
<i>FTO</i>	1.17	1.3E-12
<i>CDKAL1</i>	1.12	4.1E-11
<i>KCNJ11</i>	1.14	6.7E-11
<i>HHEX</i>	1.13	5.7E-10
<i>SLC30A8</i>	1.12	5.3E-08
<i>PPARG</i>	1.14	1.7E-06

Data from Scott et al., *Sci* 316:1343 (2007)

3 known
13 novel genes

> 67,000 samples

Gene	OR	p-value
<i>JAZF1</i>	1.10	5.0E-14
<i>CDC123/CAMK1D</i>	1.11	1.2E-10
<i>TSPAN8/LGR5</i>	1.09	1.1E-09
<i>THADA</i>	1.15	1.1E-09
<i>ADAMTS9</i>	1.09	2.1E-08
<i>NOTCH2</i>	1.13	4.1E-08

Data from Zeggini et al. *Nat Genet* 40:638, 2008

> 19,900 samples

Gene	OR	p-value
<i>KCNQ1</i>	1.4	3.1E-42

Data from Yasuda et al. *Nat. Genet.* 40:1092 (2008)

Odds Ratio (Effect Size) of variant

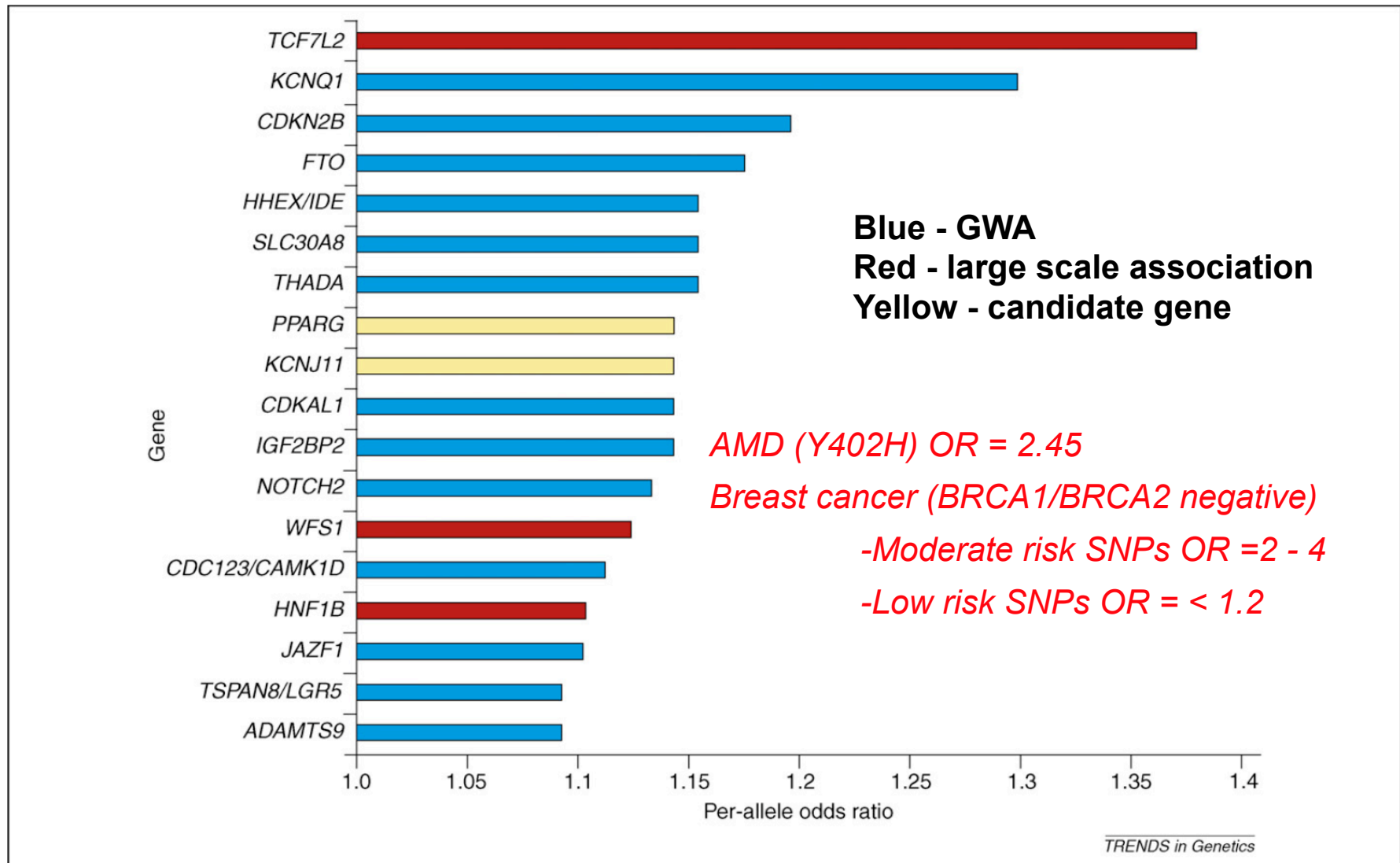
$$\text{OR} = \frac{\text{Odds of cases with or without risk allele}}{\text{Odds of controls with or without risk allele}}$$

	Have risk allele	No risk allele
case	40	60
control	25	75

$$\text{OR} = \frac{40 / 60}{25 / 75} = 2.0$$

- OR = 1** Individuals with “risk” allele are not at increased risk
- OR > 1** Implies increased risk (positive association)
- 0 < OR < 1** Implies protective effect (negative association)

Effect Sizes of the 18 Known T2D-Susceptibility Loci



TOPICS:

1. Searching for those darn genes ---GWAs

2. What have we learned?

Summary of findings

- 1. In the last 2-3 years, move from 3-18 T2D susceptibility loci**
- 2. Majority of novel variants are non-coding
(Likely not causative variants)**
 - **Aberrant gene regulation is a major factor in T2D**
 - **Challenging to tease out “mechanism”**
- 3. Low effect size and account for little of the heritability**
 - **Epidemiologically-derived λ s ~ 3.0**
 - **Combined λ s for all variants is < 1.10**

Summary of findings --con't

- 4. Additional factors need to be considered such as:**
 - **Rare variants with large effects**
 - **Other types of variation - ie. structural variation**
 - **Interactions**
 - **gene-gene**
 - **gene-environment**
 - **Gene networks (pathways)**

- 5. Beta-cell dysfunction plays a primary role in T2D**

T2D Pathophysiology

Genetic Predisposition

genetic defects



Beta-cell Dysfunction

impaired insulin
secretion

Environment

obesity



Insulin Resistance

poor glucose
utilization



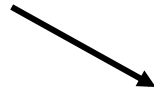
HYPERGLYCEMIA



BETA-CELL EXHAUSTION



TYPE 2 DIABETES



Mechanism of Inherited T2D-Susceptibility

Table 1. T2D-susceptibility loci for which there is genome-wide significant evidence for association^a

	Locus (nearest genes)	Year association 'proven'	Approach	Probable mechanism	Index variant	Effect size ^b	Risk-allele frequency
●	<i>PPARG</i>	2000	Candidate	Insulin action	rs1801282	1.14	0.87
★	<i>KCNJ11</i>	2003	Candidate	β-cell dysfunction	rs5215	1.14	0.35
★	<i>TCF7L2</i>	2006	Large-scale association	β-cell dysfunction	rs7901695	1.37	0.31
●	<i>FTO</i>	2007	GWA	Altered BMI	rs8050136	1.17	0.40
★	<i>HHEX/IDE</i>	2007	GWA	β-cell dysfunction	rs1111875	1.15	0.65
★	<i>SLC30A8</i>	2007	GWA	β-cell dysfunction	rs13266634	1.15	0.69
★	<i>CDKAL1</i>	2007	GWA	β-cell dysfunction	rs10946398	1.14	0.32
★	<i>CDKN2A/2B</i>	2007	GWA	β-cell dysfunction	rs10811661	1.20	0.83
● ★	<i>IGF2BP2</i>	2007	GWA	β-cell dysfunction	rs4402960	1.14	0.32
★	<i>HNF1B</i>	2007	Large-scale association	β-cell dysfunction	rs4430796	1.10	0.47
	<i>WFS1</i>	2007	Large-scale association	Unknown	rs10010131	1.12	0.60
★	<i>JAZF1</i>	2008	GWA	β-cell dysfunction	rs864745	1.10	0.50
	<i>CDC123/CAMK1D</i>	2008	GWA	Unknown	rs12779790	1.11	0.18
	<i>TSPAN8/LGR5</i>	2008	GWA	Unknown	rs7961581	1.09	0.27
	<i>THADA</i>	2008	GWA	Unknown	rs7578597	1.15	0.90
	<i>ADAMTS9</i>	2008	GWA	Unknown	rs4607103	1.09	0.76
	<i>NOTCH2</i>	2008	GWA	Unknown	rs10923931	1.13	0.10
★	<i>KCNQ1</i>	2008	GWA	β-cell dysfunction	rs2237892	1.29	0.93

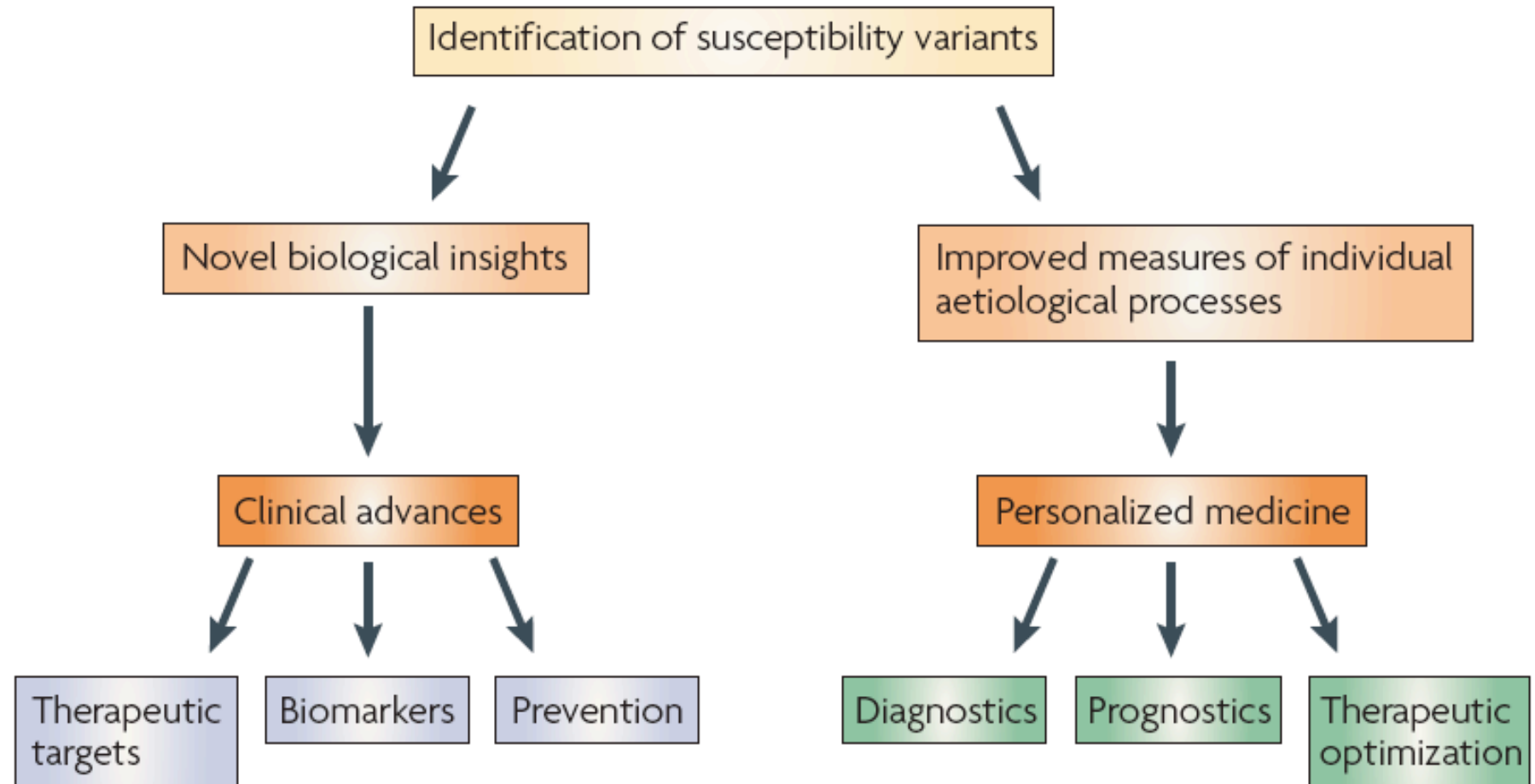
Beta-cell dysfunction plays a primary role in T2D

TOPICS:

1. Searching for those darn genes ---GWAs
2. What have we learned?
- 3. How can we use this information?**

Translate Genetic Findings to Improve Clinic Care

Box 6 | Clinical translation of findings from GWA studies

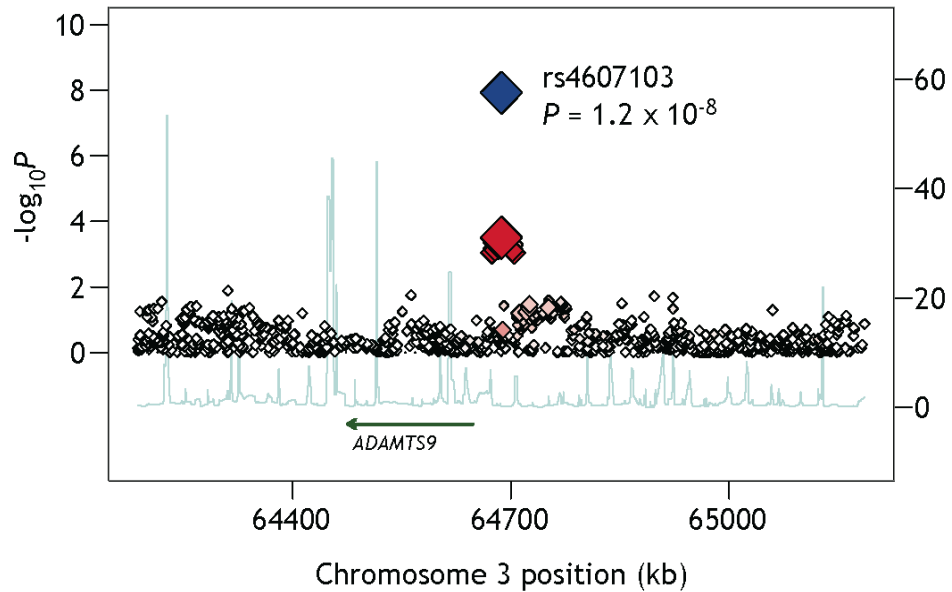


Points to Consider Regarding the Significance of GWAS Discoveries

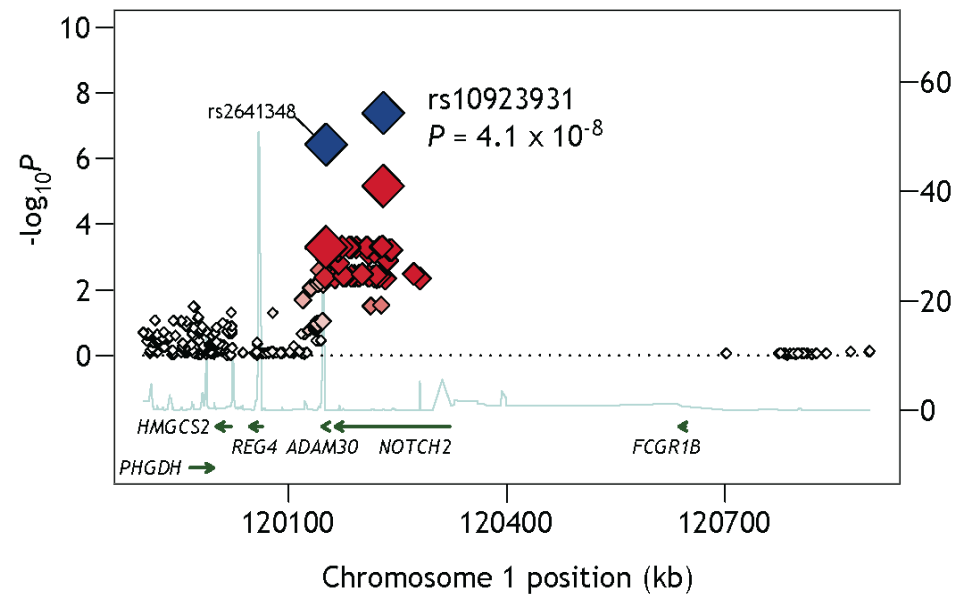
- **We assigned gene names and but we haven't identified the causative variants or genes.**

Causative Variant? Affected Gene?

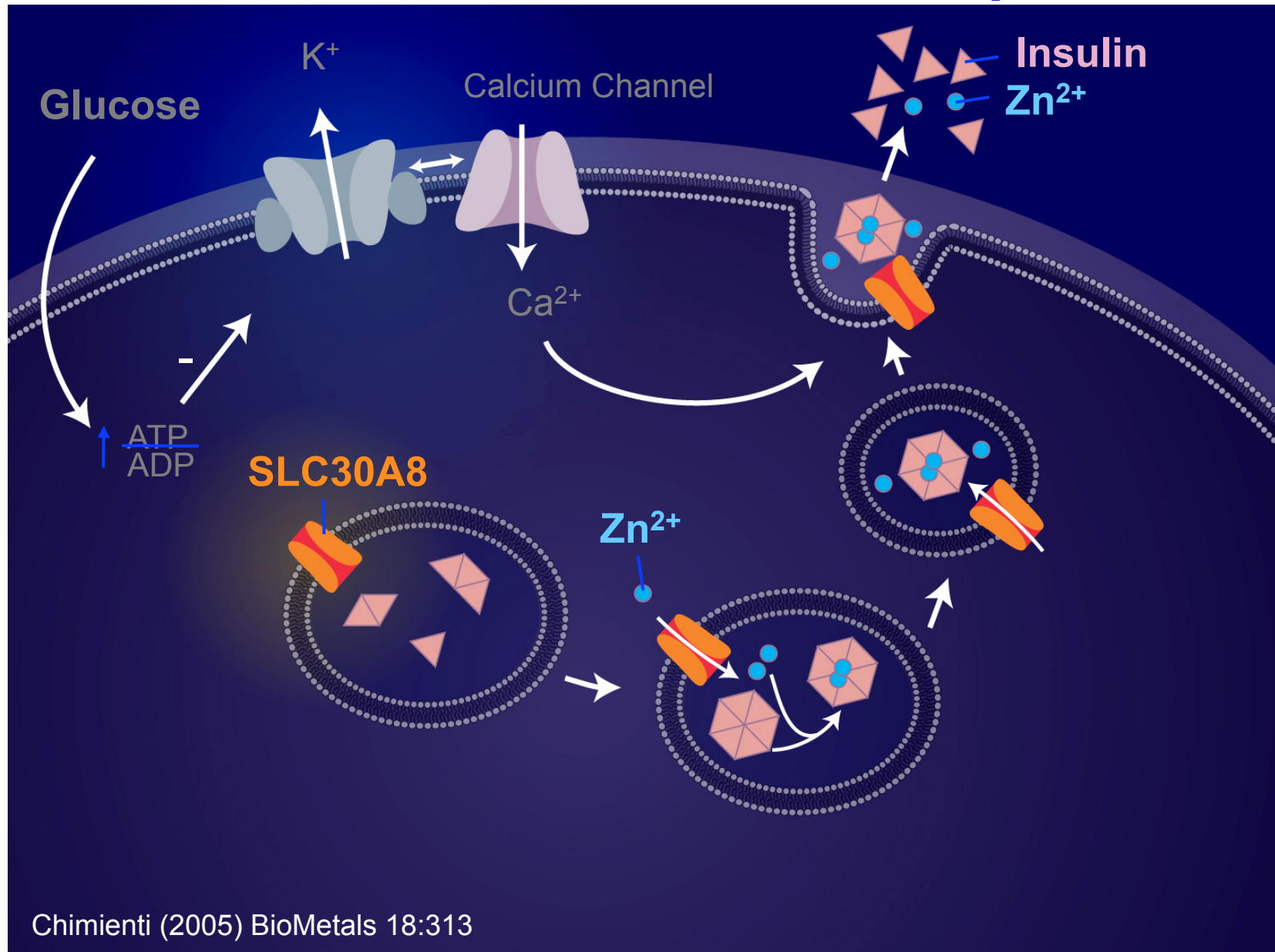
ADAMTS9 region



NOTCH2/ADAM30 region



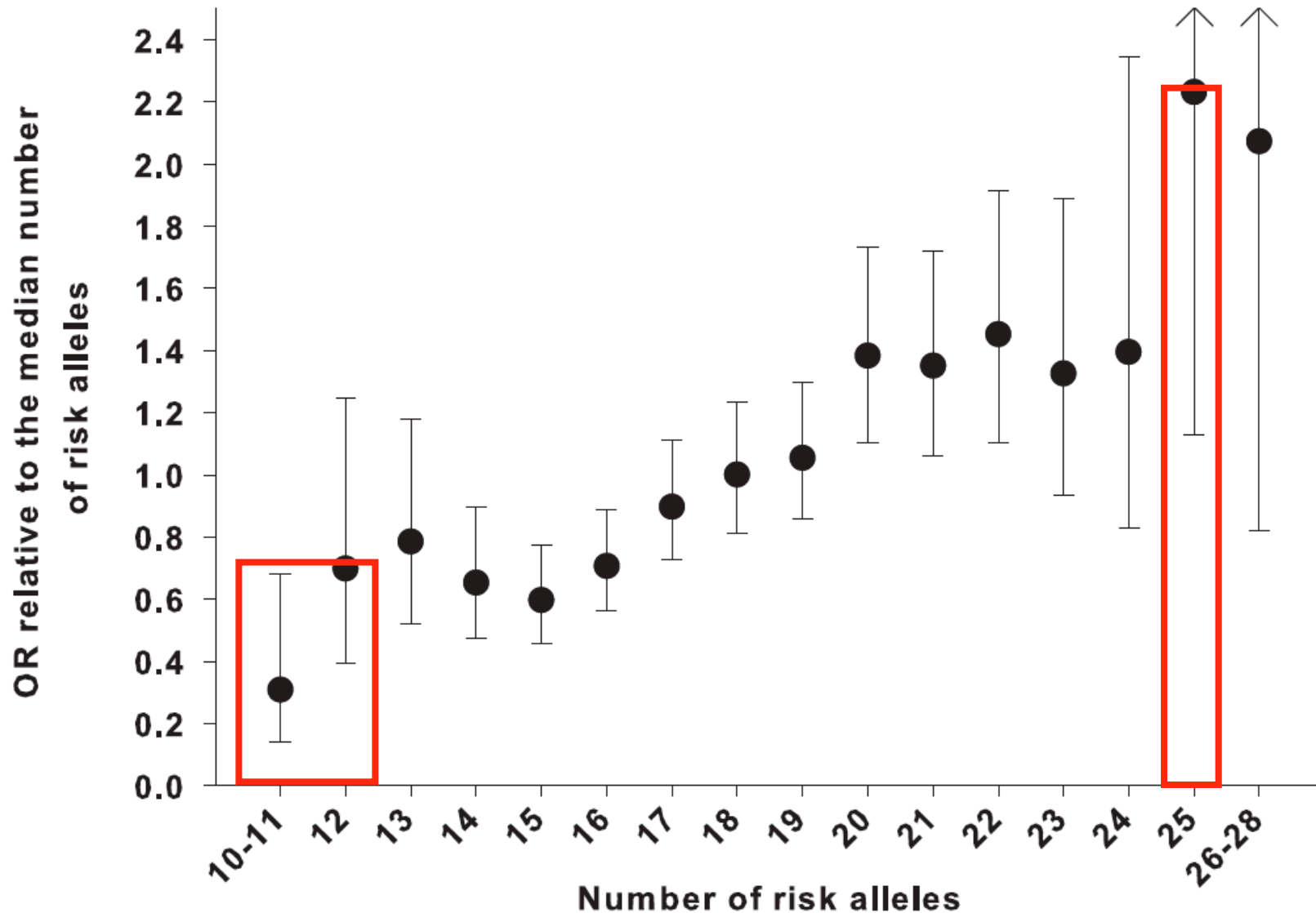
SLC30A8 - Beta Cell Zinc Transporter



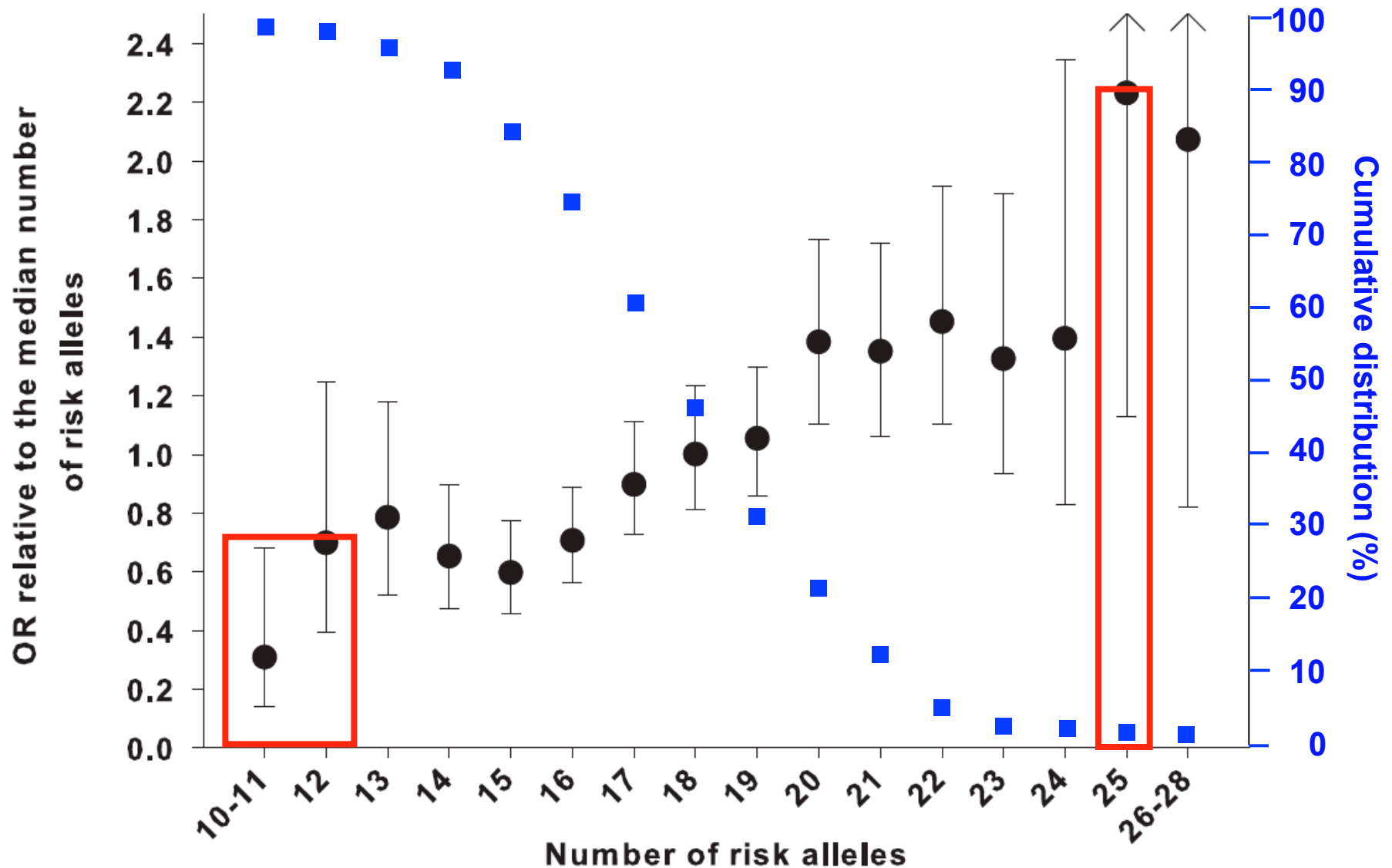
Points to Consider Regarding the Significance of GWAS Discoveries

- We assigned gene names and but we haven't identified the causative variants or genes.
- Can we applied these variants **now** to accurately predict risk of future illness in healthy people?

Increasing ORs With Increasing T2D Risk Alleles



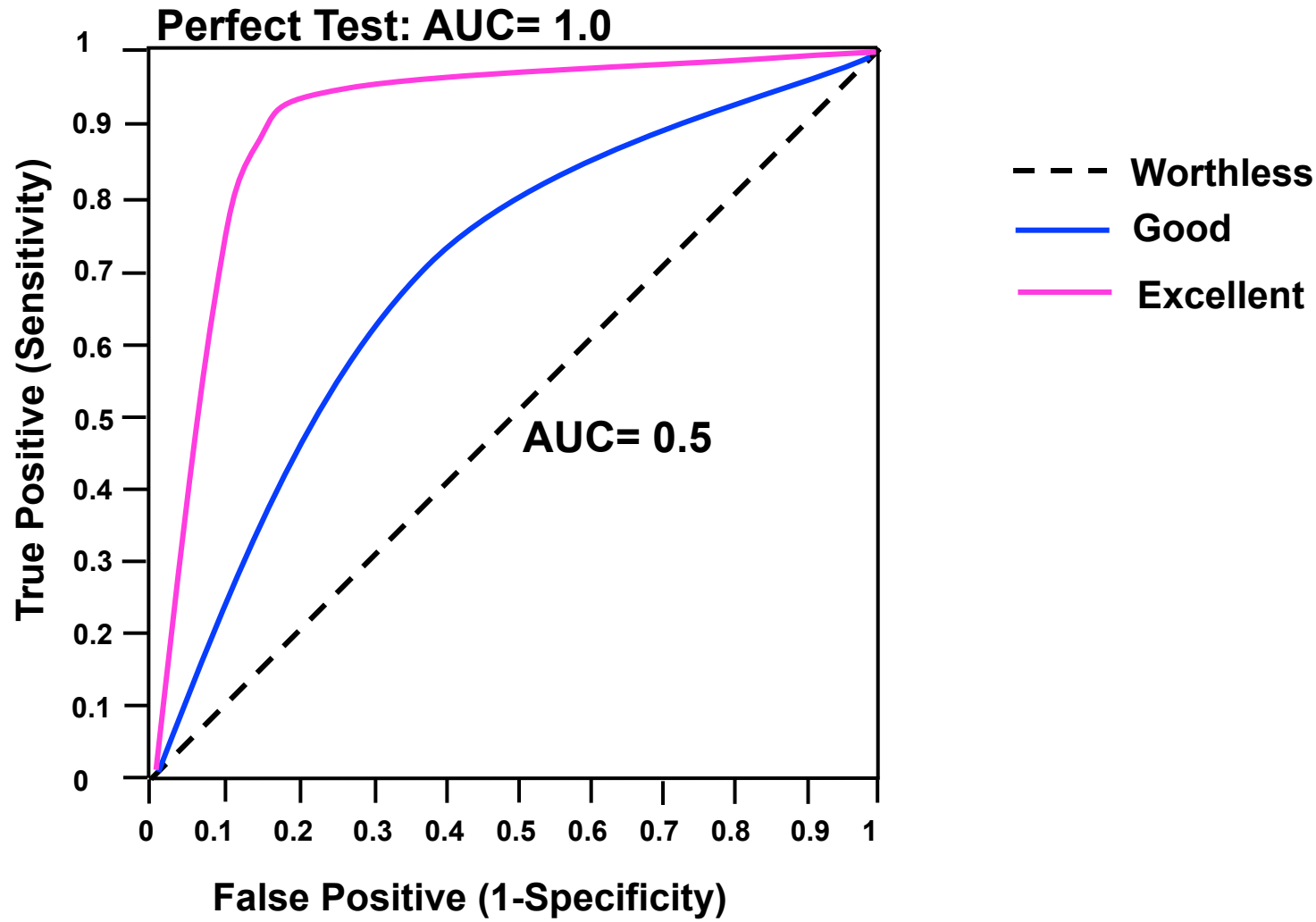
Increasing ORs With Increasing T2D Risk Alleles



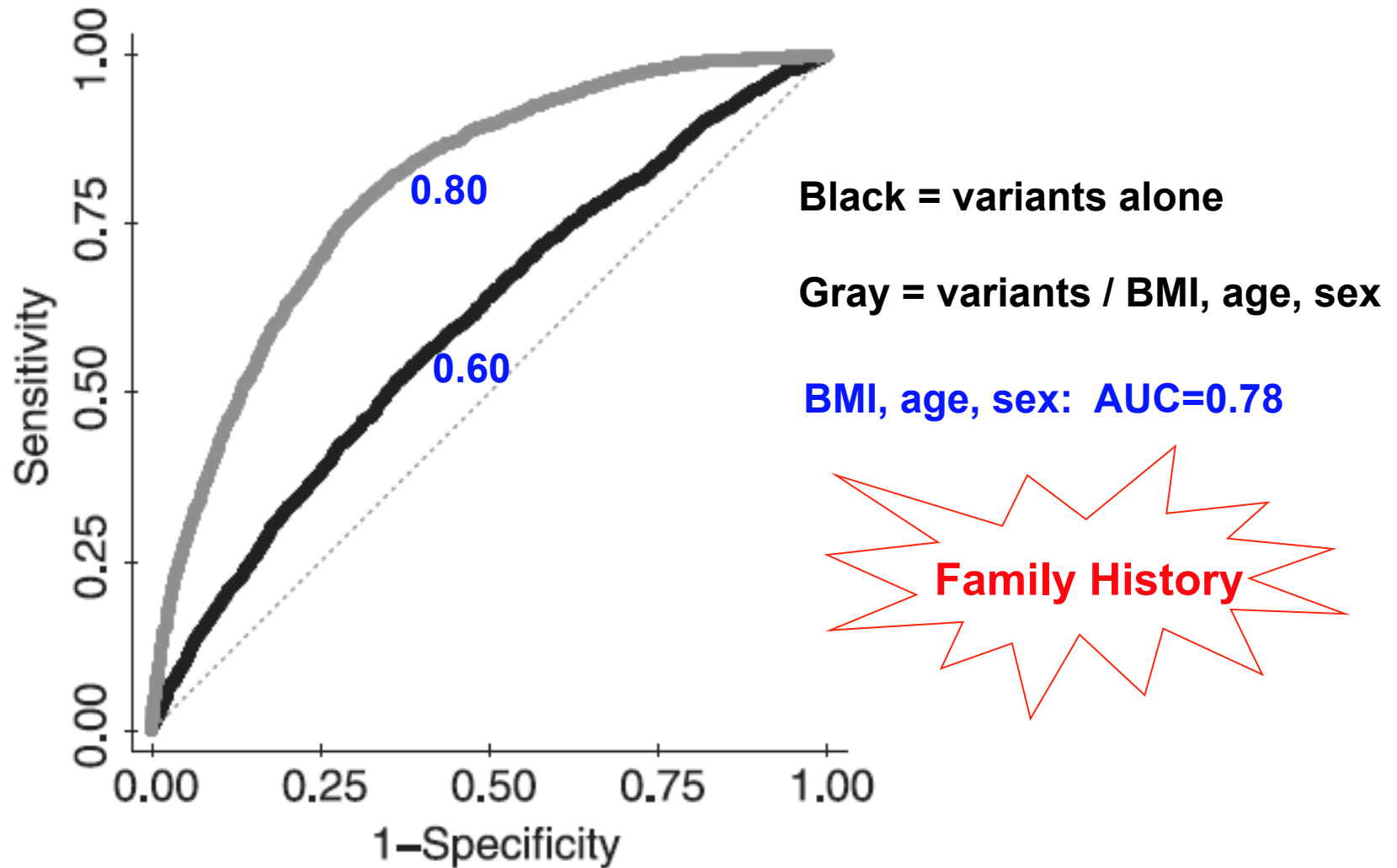
Data from Lango et al., *Diabetes* 57:3129 (2008)

Receiver Operating Characteristic Curve (ROC)

Assess the sensitivity and specificity of a test at various cutoffs.



ROC Plot of 18 T2D Risk Variants



Studies Assessing Risk Prediction scores

Ref	Study	Size	Clinical risk factors Variant	SNPs	AUC		
					Genetic risks	Clinical risks	Clinical and Genetic
Weedon (Oct-06)	UK	~6000		3	0.58		
Lyssenko (2005)	Botnia		BMI, plasma glucose level	3		0.68	0.68
Vaxillaire (Jan-08)	DESIR		Age, sex, BMI	3	0.56	0.82	0.84
Cauchi (May-08)	French	~9000		15	0.86		
Lango (Nov-08)	UK	2598	Age, sex, BMI	18	0.60	0.78	0.80
van Hoek (Nov-08)	Rotterdam	6544	Age, sex, BMI	18		0.66	0.68
Meigs (Nov-08)	Framingham Offspring	2377	Family history, , sex age, BMI, fasting glucose, SBP, HDL TG	18		0.90	0.91
Lyssenko (Nov-08)	Malmo Prevention Study and Botnia Study	~19,000	Family history, BMI, insulin secretion, insulin sensitivity, DI, BP, TG, liver enzymes, Apo AI	11	0.56-0.62	0.74	0.75
Lin (Jan-09)	CoLaus	5360	Family history, , sex age, physical activity, TG/HDL and WHR	15	0.57	0.86	0.87

Genetic testing more helpful for young individuals -absence of obvious clinical risk factors.

Genetic testing may be more predictive in the future :

- **Larger number of variants**
- **Variants with larger effect size**
- **Assess “the” causal variants**
- **Variants that correlate with “risk” traits for which we can’t assess easily via non-genetic means**

Points to Consider Regarding the Significance of GWAS Discoveries

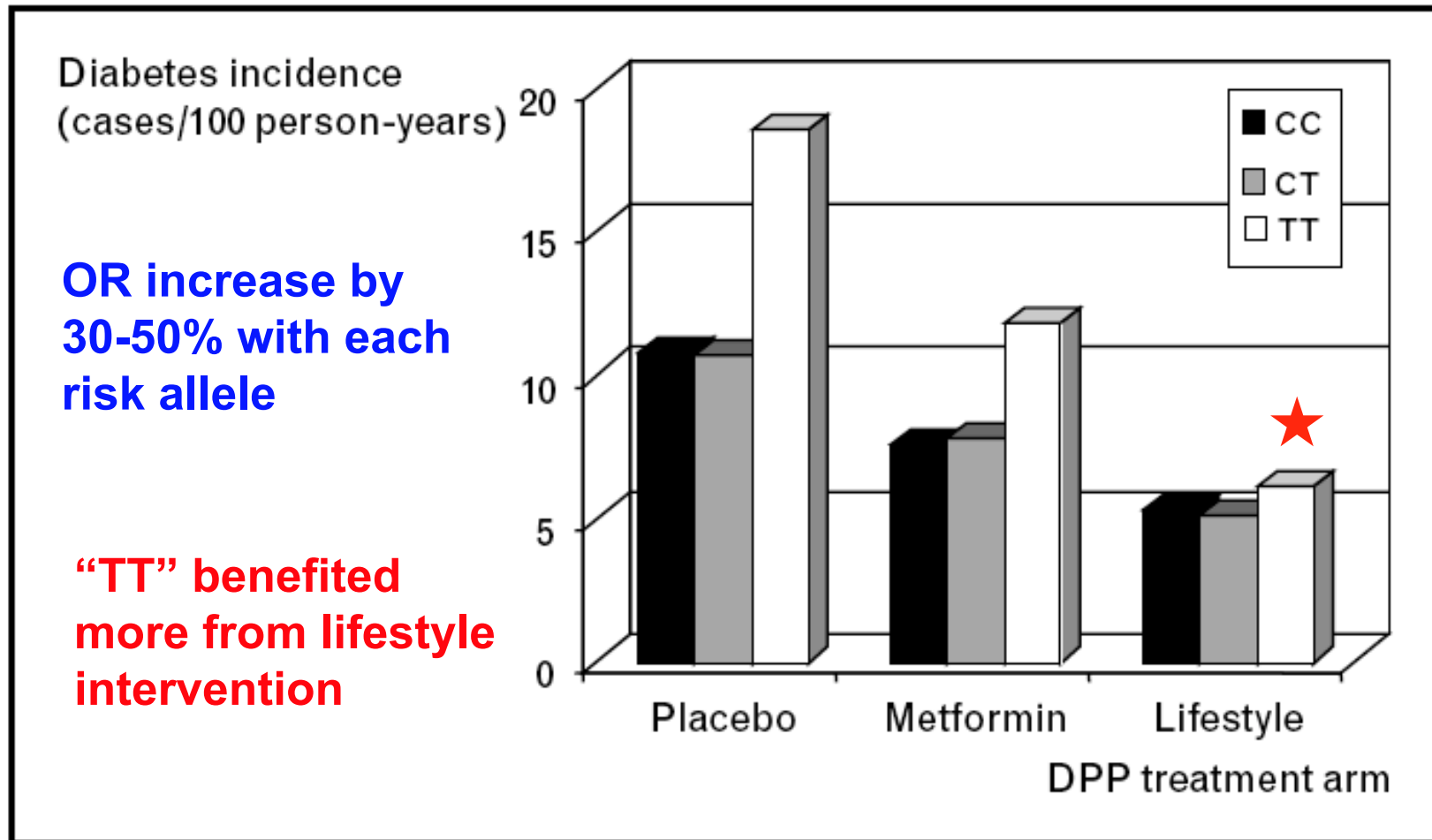
- We assigned gene names and but we haven't identified the causative variants or genes.
- Can we applied these variants now to accurately predict risk of future illness in healthy people?
- **Do we need to know the “role” of these risk factors in T2D before they can be used in personalized treatment or prevention?**

TCF7L2

- **Most replicated T2D loci and largest effect size identified to date.**
- **Associated SNPs (rs12255372 and rs7903146) are intronic --How do they affect *TCF7L2*?**
- **Transcription factor involved in Wnt: activate transcription of a variety of genes
--How does it influence diabetes risk?)**

TCF7L2 and Response to Interventions

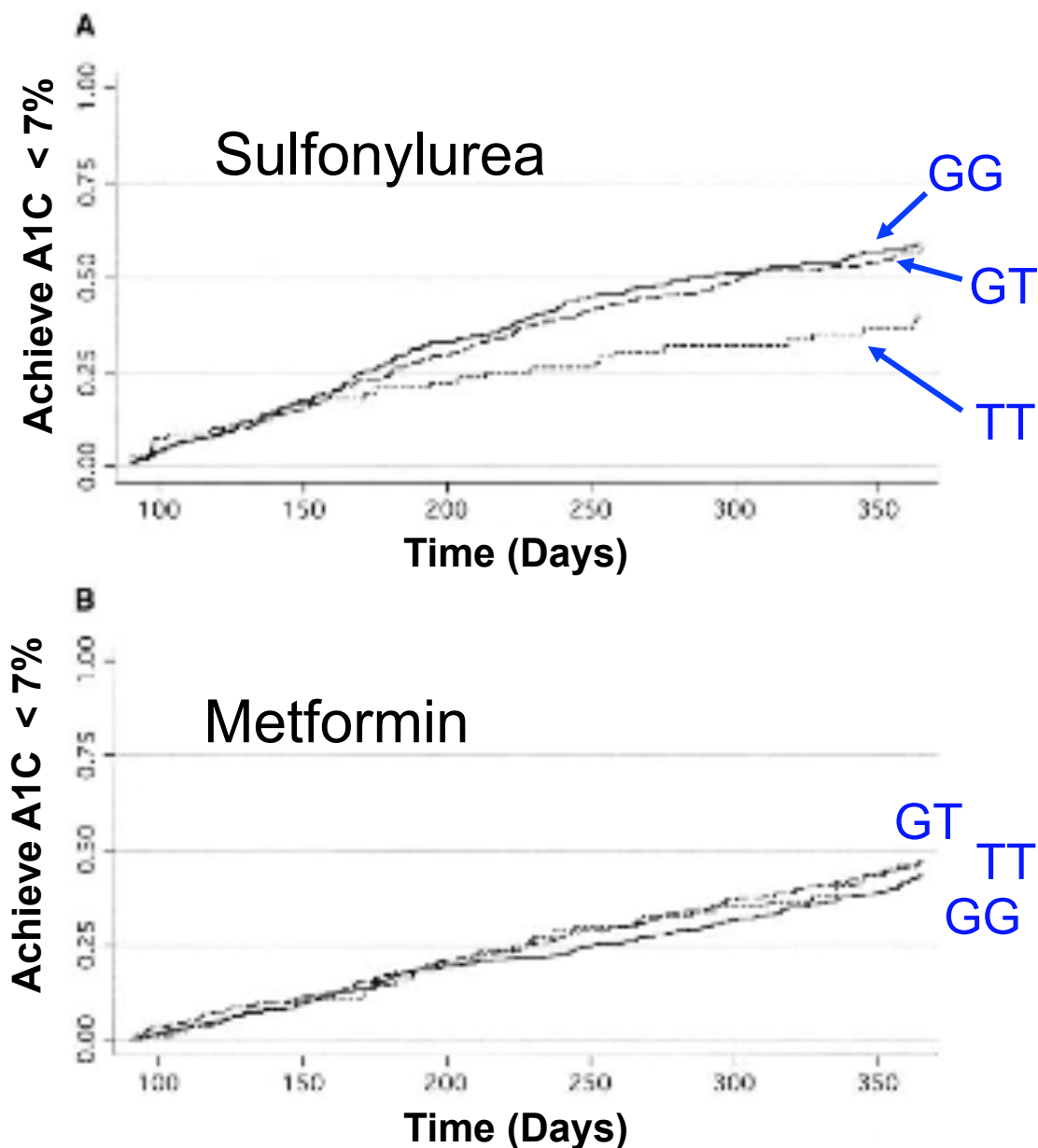
Incidence of diabetes (case/100 person-years) according to genotype at rs7903146 and treatment arm in the DPP



TCF7L2 and Response to Sulfonylurea / Metformin

GoDART Study (Scotland)

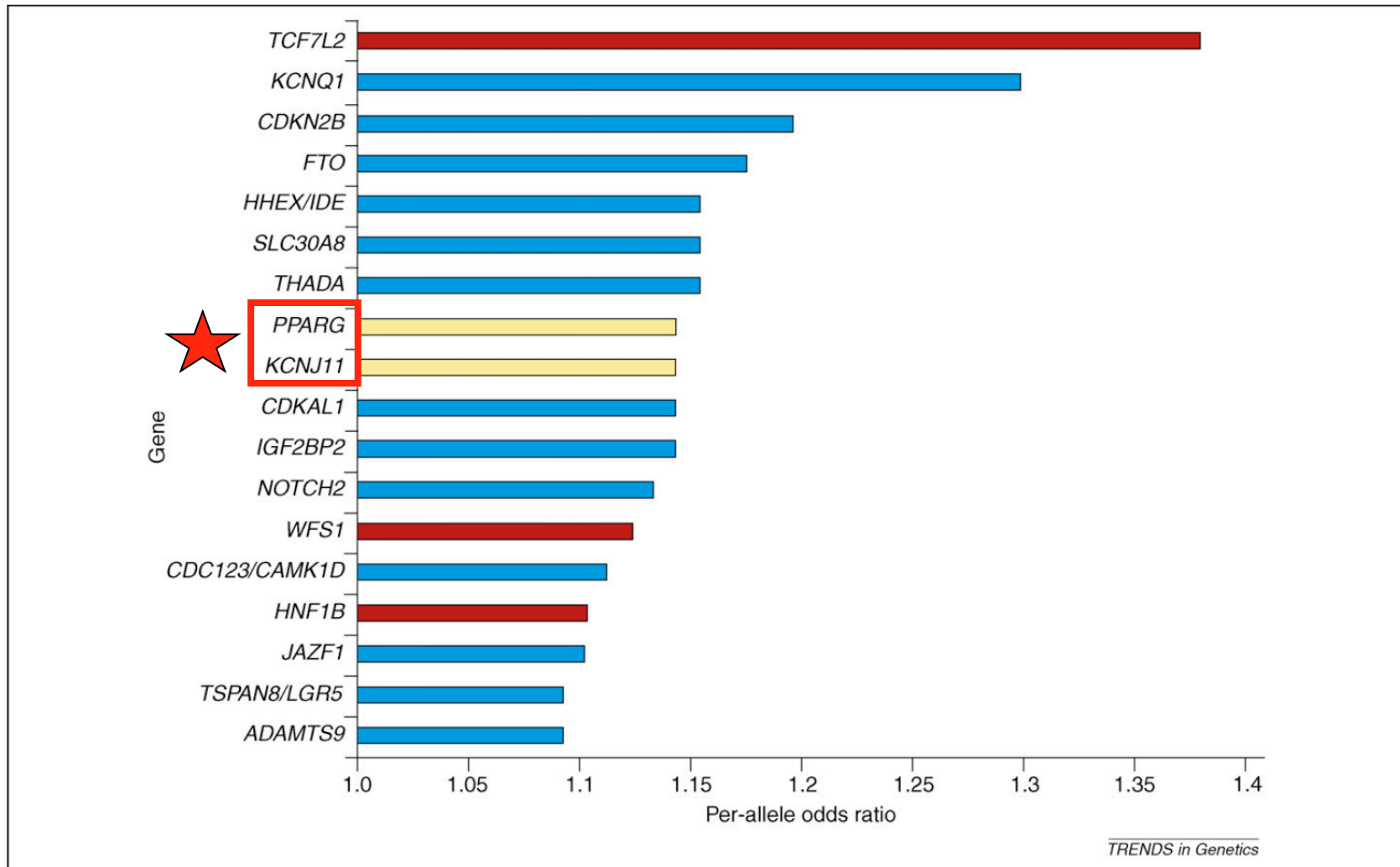
Sulfonylurea treatment failure is 2-fold greater in individuals homozygous for the rs12255372 risk allele (TT).



Points to Consider Regarding the Significance of GWAS Discoveries

- We assigned gene names and but we haven't identified the causative variants or genes.
- Can we applied these variants now to accurately predict risk of future illness in healthy people?
- Do we need to know the “role” of these risk factors in T2D before they can be used in personalized treatment or prevention?
- **Because of the small odds ratios associated with most of these variants, are they irrelevant for drug development?**

Two known major drug targets for T2D identified in the genome wide scan – aren't there likely others?



TOPICS:

1. Searching for those darn genes ---GWAs
2. What have we learned?
3. How can we use this information?
- 4. Next steps.....**

(Some) Next Steps...

1. More association studies

- **European and non European populations**
- **Utilizing other different variant types ie. CNVs**

2. Identify causal variants and mechanism of action

- **Narrow region with fine mapping**
- **Re-sequencing ---rare variants**
- **Functional studies**

3. Additional studies/analyses focused on:

- **Intermediate phenotypes**
- **Gene-Gene and Gene-Environment**
- **Pathway / Network analyses**
- **Better understanding of genome function**